

## NOTES AND INSIGHTS



**Using analytical equations to represent nonlinear relationships**Juan Ríos-Ocampo\*  and Michael Shayne Gary *Abstract*

Table functions, also referred to as graphical functions, provide a powerful and user-friendly way to represent nonlinear relationships between variables in system dynamics (SD) models. However, in many cases modelers may benefit from using analytical equations to represent nonlinear relationships for model sensitivity testing and also for communicating with researchers in other fields and disciplines. We propose six analytical equations that can be used to represent many of the nonlinear relationships commonly formulated using table functions in SD models. Specifically, this article provides guidance on using the generalized logistic function, the exponential function, the modified exponential function, the quadratic function, the logarithmic function and the power function to replace existing table functions. Importantly, we also present a version of each equation that includes an interior reference point. We demonstrate how to apply these analytical equations in SD models by replacing the table functions in the original World Dynamics model. We also provide a Python script to help implement our recommended procedure for incorporating the six analytical equations into models and a Vensim macro for each analytical equation.

Copyright © 2022 The Authors. *System Dynamics Review* published by John Wiley & Sons Ltd on behalf of System Dynamics Society.

*Syst. Dyn. Rev.* **38**, 354–370 (2022)

Additional Supporting Information may be found online in the supporting information tab for this article.

**Introduction**

Table functions have been an important aspect of system dynamics (SD) modeling since the first SD models were constructed (Forrester, 1961, 1968, 1969). Table functions enable modelers to draw on their own and other domain experts' knowledge of the system to represent nonlinear functions. Using table functions to formulate nonlinear relationships requires no advanced mathematical knowledge and provides flexibility in specifying the nature of causal relationships (Sterman, 2000). These characteristics afford distinct advantages, especially when eliciting nonlinear relationships from managers or policymakers with detailed knowledge of the system under study but who may possess limited mathematical training (e.g. see Ford and Sterman, 1998).

After following all recommended guidelines to formulate nonlinear relationships using table functions in the model conceptualization stage

UNSW Business School, University of New South Wales, Sydney, New South Wales, 2052, Australia

\* Correspondence to: Juan Ríos-Ocampo, UNSW Business School, University of New South Wales, Sydney New South Wales 2052, Australia. E-mail: [j.rios\\_ocampo@unsw.edu.au](mailto:j.rios_ocampo@unsw.edu.au)

Accepted by Hazine Rahmandad, Received 6 December 2021; Revised 20 May 2022 and 15 July 2022; Accepted 5 August 2022

System Dynamics Review

*System Dynamics Review* vol 38, No 4 (October/December 2022): 354–370

Published online in Wiley Online Library

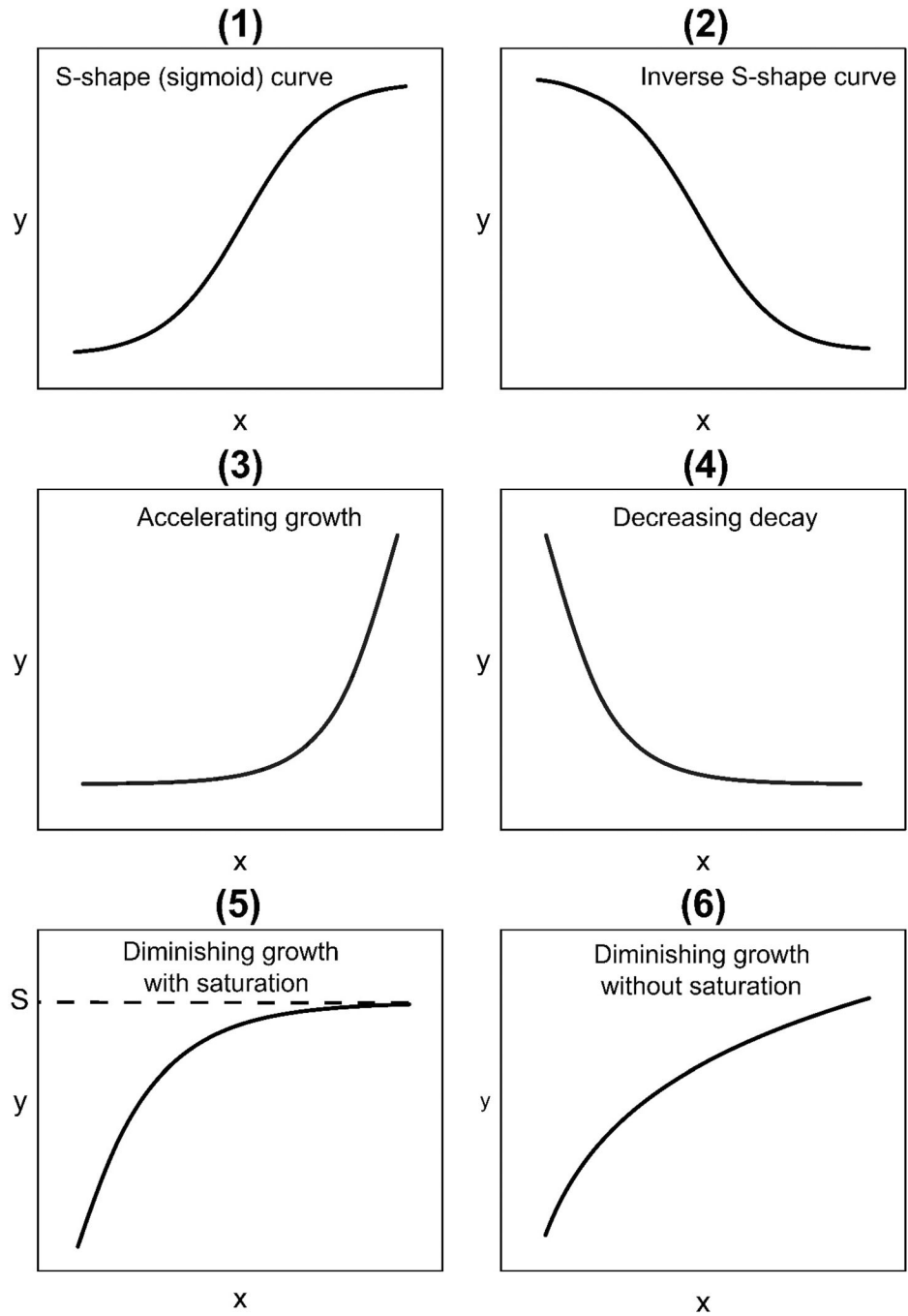
([wileyonlinelibrary.com](http://wileyonlinelibrary.com)) DOI: 10.1002/sdr.1718

(Serman, 2000: Chap. 14), continuing to use table functions when testing and communicating the model assumptions to other scholars has some disadvantages. One disadvantage involves the difficulty of running extensive sensitivity tests of different specifications for each table function in a model. Of course, modelers can test the sensitivity of model behavior to changes in one or more table function(s) by adjusting each table function manually. This process may be adequate for univariate sensitivity tests but can be time consuming when performing multivariate sensitivity tests of models with numerous table functions. Although some scholars have identified useful work-around solutions for this problem (Eker *et al.*, 2014; Gonçalves *et al.*, 2005), none of the solutions completely resolve this issue. Another disadvantage involves justifying each table function included in a model to other researchers working in the domain. Researchers unfamiliar with SD may be skeptical of table functions because nonlinear relationships in their domain may typically be represented using well-established analytical equations. An additional disadvantage of table functions includes the introduction of discontinuities from linear interpolation between points (Saeed, 1984). Increasing the number of points defined for a table function generally resolves this as a substantive concern (Eker *et al.*, 2014), but using a continuous analytical function ensures a smooth curve over the entire range of values.

Previous research proposed using analytical equations to represent nonlinear relationships that were concave, convex, or s-shaped curves (Saeed, 1984). This article builds on and extends this prior work by adding more analytical functions, so modelers can represent a larger range of nonlinear relationships. We describe six analytical equations capable of representing many of the nonlinear relationships commonly formulated using table functions in SD models. Specifically, we focus on the generalized logistic growth equation, the exponential function, the modified exponential function, the quadratic polynomial equation, the logarithmic function, and the power function. Figure 1 shows the six distinct shapes of nonlinear relationships that these analytical equations can represent: (i) S-shaped (sigmoid) curve, (ii) inverse S-shape curve, (iii) accelerating growth, (iv) decreasing decay, (v) diminishing growth with saturation, and (vi) diminishing growth without saturation. These six curves can represent a wide range of relationships and frequently appear in SD models.

We propose a procedure to convert existing table functions into analytical equations or to formulate analytical equations representing nonlinear relationships in the early model-building stages. Replacing table functions with analytical equations or adopting analytical equations directly when building an initial model facilitates more straightforward sensitivity testing of these nonlinear relationships. Also, the procedure we describe and the form of the analytical equations we use include the flexibility to specify an internal reference point for each nonlinear relationship. Providing versions of the analytical equations that include an interior reference point represents an

Fig. 1. Six distinct shapes of nonlinear relationships



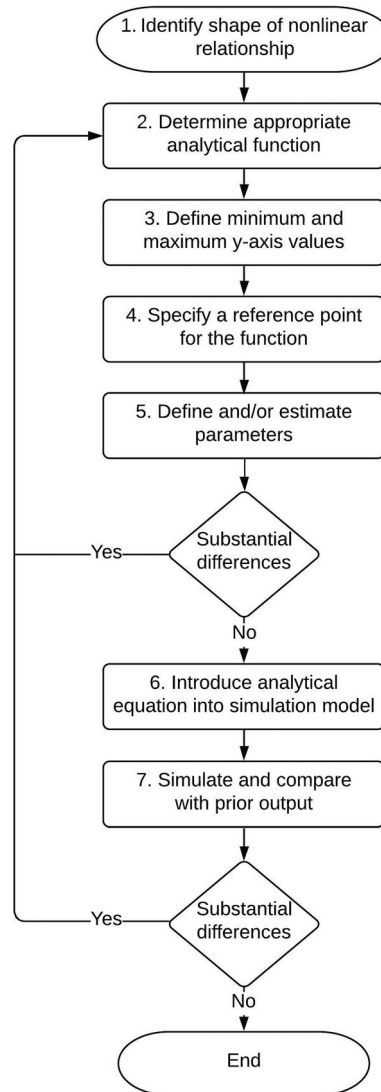
important contribution since many (perhaps most) table functions contain an interior reference point. Also, in prior research on this topic, only a small set of specific parameter conditions could be included as reference points (Saeed, 1984). In addition, providing a procedure to incorporate analytical equations into SD models may help improve communication with researchers in other fields and disciplines who frequently use analytical equations and may be skeptical of table functions.

In the next section, we outline the procedure for replacing an existing table function with one of the analytical equations. If a table function has not been formulated already, then the modeler must initially follow the guidelines for formulating nonlinear relationships (e.g. Sterman, 2000, p. 553) to determine the shape of the relationship, important reference points, and the  $(x, y)$  values for the function. The subsequent section explains the general form of each analytical equation and provides a version of each equation that includes an interior reference point  $(r_x, r_y)$ . In the final section, we provide concluding remarks. Appendices A–E, in the online supporting files, provide additional information. Appendix A provides step-by-step instructions demonstrating how to apply the procedure to replace one of the table functions in the World Dynamics model (Forrester, 1971). The World Dynamics model includes 22 table functions, and we used the same procedure to replace the other 21 table functions in the model. Due to space constraints, step-by-step instructions are only provided for replacing one table function (see the supplementary Excel file for details on replacing all 22 table functions). Appendix B explains how varying the parameters of the six analytical equations affects the shape of the functions and provides graphs illustrating the changes. Appendix C presents the steps used to derive a version of each of the analytical equations that includes an interior reference point  $(r_x, r_y)$ . Appendix D explains how to use a custom-written Python 3 script to replace an existing table function with one of the six analytical equations. Finally, Appendix E provides the syntax of a set of Vensim Macros that can be used to implement each analytical equation in Vensim.

## Procedure

This section provides the procedure for replacing table functions with analytical equations. Figure 2 illustrates the steps involved in this procedure. If a table function has not been formulated already, then Step 0 involves following the guidelines for formulating nonlinear relationships to determine the shape of the relationship, important reference points, and  $(x, y)$  values for the function (Sterman, 2000, p. 553).

Fig. 2. Procedure for using analytical equations to represent nonlinear functions



**Step 1:** Identify the shape of the nonlinear relationship in the existing table function. If the shape matches one of the curves shown in Figure 1, then one of the six analytical equations discussed in this article may be used to replace the table function.

**Step 2:** Given the shape of the nonlinear relationship, determine the analytical equation(s) (discussed further in Section 3) most appropriate for replacing the table function.

- 
- Step 3:** Define the minimum and maximum  $y$ -axis values of the curve. Modelers should adhere to good modeling practice by normalizing, where appropriate, the input and output values of the function (Sterman, 2000).
- Step 4:** Specify an important interior reference point  $(r_x, r_y)$  for the nonlinear relationship where the values of the function are determined by definition. Interior refers to a point between the minimum and maximum  $y$ -axis values. Identify a reference point, where possible, but skip this step if no reference point can be identified.
- Step 5:** If a reference point is not specified in Step 4, then use the conventional form of the analytical equation selected in Step 2. If a reference point is specified in Step 4, use the version of the analytical equation with a reference point. See Tables 1 and 2 for the conventional form of each analytical equation along with an alternative version of each equation that includes a reference point. Determine which parameters need to be defined or estimated for the selected analytical equation. Estimate the parameter values using nonlinear regression to minimize the Sum of Squared Errors (SSE) or Root-Mean-Square Error (RMSE) between the table function and the analytical equation. Large discrepancies between the original table function and the analytical equation necessitate returning to Step 2 and selecting another analytical equation or, in the case of the family of generalized logistic equations, selecting a more general form of the equation that includes additional parameters.
- Step 6:** Once the fully specified analytical equation yields a curve that closely matches the original table function, add the analytical equation to the simulation model and create new variables for parameter values, the minimum and maximum  $y$ -axis values, and the reference point.
- Step 7:** Simulate the model and compare the output for key variables before and after incorporating the analytical equations. Some differences can be expected, but large discrepancies in the before and after dynamics require further investigation and may require returning to Step 2 and selecting another analytical equation. An important point to note is that when the  $y(x)$  minimum and/or  $y(x)$  maximum values are not defined in Step 3 or do not exist, the output of the analytical equations may extend without bound and could well impact model behavior. Be sure to check that this is appropriate for the nonlinear relationship.

The above procedure can be followed to replace an existing table function with an analytical equation or to directly represent a nonlinear relationship using an analytical equation. Of course, when the modeler knows the underlying data-generating process, they can directly select the most

appropriate analytical function in Step 2 and work through the remaining steps. But modelers may not always know the underlying data-generating process. In these cases, in Step 2 modelers must carefully examine the shape of the table function curve and identify all the candidate analytical function(s) that can possibly match that qualitative shape. Next the modeler must go through the remaining steps of our procedure for each candidate analytical equation and then use all the results combined with judgment to select the most appropriate analytical equation to replace the table function. The choice of which analytical function to use must not be based only on the best statistical fit.

We have also written a Python script—provided in the online supporting information and explained in Appendix D—that helps automate this procedure. When using the script, the modeler enters information about the existing table function(s), and the program estimates the best-fitting values for each of the six analytical equations based on the SSE statistics and RMSE. The script also draws a graph showing all the estimated curves along with the original table function. The modeler still must use all the results combined with judgment to select the most appropriate analytical equation to replace the table function.

## Analytical equations

This section introduces a set of exact, closed-form analytical equations that can be used to represent a range of different nonlinear relationships. We focus on: (i) the generalized logistic function, (ii) the exponential function, (iii) the modified exponential function, (iv) the quadratic function, (v) the logarithmic function, and (vi) the power function. After explaining the conventional form of each analytical equation, we also present a version of each equation that includes a reference point,  $(r_x, r_y)$ , where the values of the function are determined by definition. For example, many nonlinear functions in SD models have  $(1, 1)$  as a reference point, indicating the relationship between the two variables under normal conditions (Forrester, 1971; Martin, 1997; Sterman, 2000).

### *Generalized logistic function*

The generalized logistic function is used in many disciplines, including the natural sciences and social sciences. The generalized logistic function, provided in Eq. (1), can represent nonlinear relationships characterized by an S-shape or sigmoid curve (Graph 1 in Figure 1), the inverse S-shape curve (Graph 2 in Figure 1), as well as the diminishing growth with saturation curve (Graph 5 in Figure 1). Using this equation involves identifying the minimum value of  $y(x)$ , the inflection point, and the maximum value of

$y(x)$ . This function is widely used in modeling a variety of different phenomena such as cumulative sales in marketing (Lilien *et al.*, 2017), technology adoption (Struben and Sterman, 2008), and the spread of infections in pandemics (Struben, 2020).

$$y(x) = A + \frac{K - A}{(C + Q * e^{-\beta * (x - M)})^{\frac{1}{\nu}}} \quad (1)$$

where  $A$ ,  $K$ ,  $C$ ,  $\beta$ ,  $Q$ ,  $M$ , and  $\nu$  are parameters and  $x$  is the independent variable.  $A$  is the minimum value (or lower asymptote) of  $y(x)$ . When  $C = 1$ ,  $K$  is the maximum value (or upper asymptote) of  $y(x)$ . If  $A = 0$  and  $C = 1$ ,  $K$  is known as the carrying capacity. The parameter  $M$  represents the inflection point of the curve. When  $\beta > 0$  the curve yields S-shaped growth and when  $\beta < 0$  the curve yields the inverse (or reverse) S-shaped decline. Appendix B in the online supporting information explains how varying each parameter affects the logistic curve shape and provides graphs illustrating the changes. Please note that throughout the article and online supporting information, Greek letters ( $\beta$ ,  $\theta$ , and  $\lambda$ ) denote growth rates and Roman letters are used to denote all other parameters.

The generalized logistic equation can also be simplified depending on the particular shape of the nonlinear relationship the modeler wants to represent. These simplified versions compose alternative specifications of the generalized logistic function. For example, in one of the most common specifications (provided in Eq. (2)):  $C$ ,  $Q$ , and  $\nu = 1$ ,  $A$  is the minimum value of  $y(x)$ , and  $K$  is the maximum value of  $y(x)$ .

$$y(x) = A + \frac{K - A}{1 + e^{-\beta * (x - M)}} \quad (2)$$

Table 1 provides a number of alternative specifications for the generalized logistic equation alongside an equivalent version of each equation that includes a reference point  $(r_x, r_y)$ . We recommend starting with the most common, and simpler, version of the generalized logistic function in which the parameters  $C$ ,  $Q$  and  $\nu = 1$  (see Appendix B in the online supporting information). If this simpler form cannot adequately match the shape of the focal table function, then choose an alternative specification with more degrees of freedom.

We provide a version of each equation that includes an interior reference point—in addition to the minimum and maximum  $y(x)$  values—because many table functions include an interior reference point. For example, many table functions include  $(1, 1)$  as a reference point. The version of the analytical equations that include a reference point ensures modelers can specify and constrain the function to pass through that reference point. To



Table 1. Alternative generalized logistic-equation specifications and equivalent versions with a reference point

Generalized logistic equation		
Specification	Conventional equation	Version of equation with reference point
Most general form	$y(x) = A + \frac{K-A}{(C+Qe^{-\beta(x-M)})^v}$	None provided
$v = 1$	$y(x) = A + \frac{K-A}{C+Qe^{-\beta(x-M)}}$	$y(x) = A + \frac{K-A}{C+Q* \left( \frac{A*(C-1)+K-r_y*C}{Q*(r_y-A)} \right)^{\frac{x-M}{r_x-M}}}$
$v, Q, C = 1$	$y(x) = A + \frac{K-A}{1+e^{-\beta(x-M)}}$	$y(x) = A + \frac{K-A}{1+ \left( \frac{K-r_y}{r_y-A} \right)^{\frac{x-M}{r_x-M}}}$
$v, Q, C = 1; A = 0$	$y(x) = \frac{K}{1+e^{-\beta(x-M)}}$	$y(x) = \frac{K}{1+ \left( \frac{K-r_y}{r_y} \right)^{\frac{x-M}{r_x-M}}}$

incorporate a reference point  $(r_x, r_y)$  into the equation, we rearranged terms in Eq. (1), solved for  $\beta$  assuming  $v = 1$ , and then inserted that solution back into the original equation. See Appendix C in the online supporting information for derivations of the analytical equations with a reference point. The slope of the curve is negative when  $r_x > M$  and  $r_y < M$  or  $r_x < M$  and  $r_y > M$  and is positive when  $r_x, r_y > M$  or  $r_x, r_y < M$ . For example, Eq. (3) provides the generalized logistic function with a reference point  $(r_x, r_y)$  when  $v = 1$ .

$$y(x) = A + \frac{K - A}{C + Q * \left( \frac{A*(C-1)+K-r_y*C}{Q*(r_y-A)} \right)^{\frac{x-M}{r_x-M}}} \tag{3}$$

With the following constraints:

$$\begin{aligned} K &> r_y > A \\ r_x &\neq M \\ Q * (r_y - A) &\neq 0 \end{aligned}$$

*Exponential function*

The exponential function, provided in Eq. (4), is well-known and widely used to represent growth or decay dynamics in many disciplines (Al-Eideh and Al-Omar, 2019; Dennis, 1989). The exponential function can represent accelerating growth (Graph 3 in Figure 1) or decreasing decay (Graph 4 in Figure 1).

$$y(x) = d * e^{\theta * x} + c \tag{4}$$

where  $x$  is the independent variable and  $\theta$ ,  $d$ , and  $c$  are parameters.  $\theta$  represents the intrinsic growth rate:  $\theta > 0$  yields exponential growth and  $\theta < 0$  yields exponential decay. Appendix B in the online supporting information explains how varying each parameter affects the shape of the exponential curve and provides graphs illustrating the changes.

To incorporate a reference point  $(r_x, r_y)$  into the equation, we rearranged terms in Eq. (4), solved for  $\theta$ , and then inserted that solution back into the original equation. Equation 5 provides the exponential function with a reference point  $(r_x, r_y)$ . See Appendix B in the online supporting information for the derivation.

$$y(x) = d * \left( \frac{r_y - c}{d} \right)^{\frac{x}{r_x}} + c \tag{5}$$

where the following constraints apply:

$$\begin{aligned} \frac{r_y - C}{d} &> 0 \\ r_x &\neq 0 \\ r_y &> C \\ d &\neq 0 \end{aligned}$$

### Modified exponential function

The modified exponential function, provided in Eq. (6), can represent growth dynamics with decreasing returns to scale that reaches a saturation level (Graph 5 in Figure 1).

$$y(x) = a * (1 - e^{-\lambda * x}) + w \tag{6}$$

where  $x$  represents the independent variable and  $\lambda$ ,  $w$ , and  $a$  are parameters.  $x$  must be positive. When  $x = 0$ ,  $w$  is the minimum value of  $y(x)$ . When  $w = 0$ ,  $a$  is the maximum value of  $y(x)$  or the saturation level. When  $w$  does not equal 0,  $a + w$  is the maximum value of  $y(x)$ . Appendix B in the online supporting information explains how varying each parameter affects the shape of the modified exponential curve and provides graphs illustrating the changes.

To incorporate a reference point  $(r_x, r_y)$  into the equation, we rearranged terms in Eq. (6), solved for  $\lambda$ , and then inserted that solution back into the original equation. Equation (7) provides the modified exponential function

with a reference point  $(r_x, r_y)$ . See Appendix C in the online supporting information for the derivation.

$$y(x) = a * \left( 1 - \left( \frac{a + w - r_y}{a} \right)^{\frac{x}{r_x}} \right) + w \quad (7)$$

where the following constraints apply:

$$\begin{aligned} a &\neq 0 \\ r_x &\neq 0 \\ r_y &< a + w \end{aligned}$$

### Quadratic function

The quadratic function, shown in Eq. (8), is a polynomial of degree two that can represent accelerating growth that is not exponential (i.e. Graph 3 in Figure 1 for concave growth curves that are not exponential). However, care must be taken when using the quadratic function because the shape of the Quadratic function can become nonmonotonic when one or more of the parameters are negative.

$$y(x) = a + b * x + c * x^2 \quad (8)$$

where  $a$ ,  $b$  and  $c$  are parameters.  $a$  defines the y-axis intercept when  $y(x=0)$ . If  $c=0$ , the equation is linear. The  $c$  parameter controls the degree of curvature of the graph; a larger magnitude of  $c$  gives the graph a more closed (sharply curved) appearance. Appendix B in the online supporting information contains graphs illustrating how varying  $c$  affects the shape of the quadratic function.

To incorporate a reference point  $(r_x, r_y)$  into the equation, we rearranged terms in Eq. (8), solved for  $c$ , and then inserted that solution back into the original equation. Equation (9) provides the quadratic equation with a reference point  $(r_x, r_y)$ . See Appendix C in the online supporting information for the derivation.

$$y(x) = a + b * x + \left( \frac{r_y - a - b * r_x}{r_x^2} \right) * x^2 \quad (9)$$

where the  $r_x \neq 0$  constraint applies.

### Logarithmic function

The logarithmic function, shown in Eq. (10), can represent nonlinear relationships where growth or decay is fast at the beginning and slows down over time. This function exhibits decreasing returns to scale but continues growing (Graph 6 in Figure 1) or decaying and does not saturate or reach a limit.

$$y(x) = a + b * \ln(x) \quad (10)$$

where  $x$  ranges from 0 to  $\infty$ , and  $a$  and  $b$  are parameters. Changes in  $a$  move the start of the curve up or down along the  $y$ -axis. If  $b > 0$  the curve increases with diminishing returns and if  $b < 0$  the curve decreases with diminishing returns. Appendix B in the online supporting information contains graphs illustrating how varying  $b$  affects the shape of the logarithmic curve.

To incorporate a reference point  $(r_x, r_y)$  into the equation, we rearranged terms in Eq. (10), solved for  $b$ , and then inserted that solution back into the original equation. Equation 11 provides the logarithmic equation with a reference point  $(r_x, r_y)$ . See Appendix C in the online supporting information for the derivation.

$$y(x) = a + \left( \frac{r_y - a}{\ln(r_x)} \right) * \ln(x) \quad (11)$$

When  $r_y > a$ , the curve increases, and when  $r_y < a$ , the curve decreases, and where the following constraints apply:

$$\begin{aligned} r_x &> 1 \\ x &> 0 \end{aligned}$$

### Power function

The power function, shown in Eq. (12), can represent accelerating growth, diminishing growth with and without saturation (Graphs 3 and 6 in Figure 1), and also linear growth.

$$y(x) = a * x^\beta + c \quad (12)$$

where  $a$ ,  $\beta$  and  $c$  are parameters. Changes in  $c$  move the start of the curve up or down along the  $y$ -axis, and changes in  $a$  alter the height of the curve vertically. If  $\beta = 1$ , the equation is linear. If  $x > 0$  and  $\beta > 1$ , the shape of the function is accelerating growth (Graph 3 in Figure 1). If  $\beta < 1$ , the shape of the function is growth with decreasing returns (but continues growing; see

Graph 6 in Figure 1). Also, if  $\beta \leq 0$ , then the following constraint applies:  $x \neq 0$ . Appendix B in the online supporting information contains graphs illustrating how varying the parameters affects the shape of the power function.

To incorporate a reference point  $(r_x, r_y)$  into the equation, we rearranged terms in Eq. (12), solved for  $\beta$ , and then inserted that solution back into the original equation (Eq. (12)). Equation (13) shows the power-function equation with a reference point  $(r_x, r_y)$ . See Appendix C in the online supporting information for the derivation.

$$y(x) = a * x^{\left(\frac{\ln\left(\frac{r_y - c}{a}\right)}{\ln(r_x)}\right)} + c \tag{13}$$

where the following constraints apply:

$$\begin{aligned} r_x &> 1 \\ r_y &> c \\ a &> 0 \end{aligned}$$

Table 2 provides a summary of the conventional forms of the exponential, modified exponential, quadratic, logarithmic, and power equations along with equivalent versions that include an interior reference point.

In Appendix A in the online supporting information, we demonstrate our procedure for replacing table functions with analytical equations using the World Dynamics model (Forrester, 1971). We start by selecting one table function for a step-by-step demonstration of the seven steps. Subsequently, we provide a summary table of the analytical equations we use to replace the remaining 21 table functions in the model. Then, we compare the model behavior between the original table-function formulations and the version of the model with analytical

Table 2. Conventional form of exponential, modified exponential, quadratic, logarithmic, and power functions with equivalent versions with a reference point

	Conventional equation	Version of equation with reference point
Exponential function	$y(x) = d * e^{\theta * x} + c$	$y(x) = d * \left(\frac{r_y - c}{d}\right)^{\frac{x}{r_x}} + c$
Modified exponential function	$y(x) = a * (1 - e^{-\lambda * x}) + w$	$y(x) = a * \left(1 - \left(\frac{a + w - r_y}{a}\right)^{\frac{x}{r_x}}\right) + w$
Quadratic function	$y(x) = a + b * x + c * x^2$	$y(x) = a + b * x + \left(\frac{r_y - a - b * r_x}{r_x^2}\right) * x^2$
Logarithmic function	$y(x) = a + b * \ln(x)$	$y(x) = a + \left(\frac{r_y - a}{\ln(r_x)}\right) * \ln(x)$
Power function	$y(x) = a * x^\beta + c$	$y(x) = a * x^{\left(\frac{\ln\left(\frac{r_y - c}{a}\right)}{\ln(r_x)}\right)} + c$

equations. In the online supporting information, we include an Excel file showing the analytical equation used to replace each of the 22 table functions in the original World Dynamics model. Each tab shows a different table function, the replacement analytical equation, and the estimated fit between the two.

In addition, in Appendix E of the online supporting information, we include the syntax for Vensim Macros that can be used to incorporate each of the conventional analytical equations or the version of the equations with an interior reference point. We also add the Macros files to the online supporting information to implement in Vensim DSS and a Vensim model for each analytical equation demonstrating how to use each Macro.

The Python script and the Vensim DSS Macro files can be downloaded from <https://github.com/jprio2/Using-Analytical-Equations-to-Represent-Nonlinear-Relationships>.

## Discussion and conclusion

This article proposes a procedure for using a set of analytical equations to replace table functions or to formulate nonlinear relationships using analytical equations directly when building an initial model for representing some of the most common nonlinear relationships in SD models. To facilitate using these analytical equations, we provide a version of the generalized logistic function, the exponential function, the modified exponential function, the quadratic function, the logarithmic function and the power function that includes an interior reference point. We explain each of these analytical functions and the differently shaped nonlinear relationships each function can represent. In Appendix A in the online supporting information, we also demonstrate how to replace one table function from the original World Dynamics model with an analytical equation, and we compare the behavior of the original World Dynamics model with a version of the model in which all 22 table functions were replaced with analytical equations.

While table functions provide a powerful and user-friendly way to represent nonlinear relationships between variables in SD models, in some cases modelers will benefit from using analytical equations to replace table-function formulations. In particular, analytical equations are continuous functions, unlike piecewise linear table functions. Replacing table functions with analytical equations removes any discontinuities arising from linear interpolation between points and also ensures correct calculation of values beyond the defined range of input values. Also, analytical equations can improve communication with researchers in other fields and disciplines who typically represent nonlinear relationships using analytical equations.

In addition, using analytical equations can simplify multivariate model sensitivity testing when a model contains many nonlinear relationships. Prior research has identified useful work-around solutions for multivariate

sensitivity testing using table functions (Eker *et al.*, 2014; Gonçalves *et al.*, 2005), but none of the solutions completely resolve this issue. Eker *et al.* (2014) illustrate an approach that involves multiplying each table function by a distortion function for testing the sensitivity of the model to variations in the table functions. The parameters of the distortion function can be varied as part of subsequent sensitivity testing. While this process does enable multivariate sensitivity testing of models with table functions, selecting appropriate distortion functions and ranges of parameter values can be time consuming and involve significant trial and error.

Gonçalves *et al.* (2005) adopt a different approach that involves defining two different table functions for each nonlinear relationship and specifying a weighting parameter to compute the weighted average table function. Sensitivity testing includes varying the weighting parameter to understand how model behavior changes in response to different values of the weighted average function. This approach could be readily extended to accommodate  $N$  different table functions for each nonlinear relationship. Once again, this approach enables multivariate sensitivity testing of models with table functions, but it does not remove any discontinuities arising from linear interpolation between points in the table functions and may be challenging to explain to researchers in other fields and disciplines who typically represent nonlinear relationships using analytical equations.

At the same time, we strongly advocate that modelers use table functions to initially formulate nonlinear relationships by following the guidelines for good modeling practice (Sterman, 2000: Chap. 14). This increases the likelihood of ending up with well-formulated nonlinear relationships that can be replaced with analytical equations at a later stage. Table functions also still provide more flexibility in terms of representing nonlinear relationships than analytical equations. For example, nonlinear relationships containing more than one interior reference point may not be accurately represented using the specifications of the analytical equations presented in this article. In addition, table functions afford a more user-friendly approach for exploring alternative formulations for nonlinear relationships when experimenting with the model alongside domain experts or clients. Clients typically possess deep knowledge about the systems in which they operate, but they may not understand the analytical equations. In such cases, continued use of the original table functions helps maintain transparency of model assumptions and empowers clients to “own” the model. Modelers can plan to replace table functions with analytical equations later in the model-development process for sensitivity testing and communicating with other scholars.

We believe this article will help modelers make use of analytical equations to replace table functions or to help formulate nonlinear relationships using analytical equations directly when building an initial model. Many published SD models contain analytical functions, but there has been limited guidance (an exception is Saeed, 1984) available for modelers who may not feel

confident working with analytical equations. This article builds on and extends this prior work (Saeed, 1984) by adding more analytical equations that can be used to represent a larger range of nonlinear relationships. In addition, the procedure proposed in this article enables modelers to incorporate an interior reference point when using the six analytical equations we discussed. To enable modelers to incorporate a reference point, we derived versions of each of the analytical equations that explicitly include a reference point. However, as mentioned above, one limitation of our derivations is that each analytical equation can contain only one reference point within the interior of the curve (i.e. apart from the minimum and maximum values of the function). Table functions that contain multiple interior reference points may not be appropriate candidates for replacement using the analytical equations we derived. A promising opportunity exists for future research to extend our work to include multiple interior reference points for the six analytical functions we discussed. In addition, future research that further expands the set of analytical equations we discussed would make an important contribution.

### Acknowledgement

We thank Jose Ríos-Ocampo and Carlos Ríos-Ocampo for their contributions to the Python script and the mathematical derivations. We also thank Managing Editor Hazhir Rahmandad and two anonymous reviewers for their constructive comments and valuable suggestions to improve the article. Finally, the first author thanks the UNSW Scientia PhD Scholarship program. [Correction added on 19 September 2022, after first online publication: Acknowledgement section has been updated in this version.] Open access publishing facilitated by University of New South Wales, as part of the Wiley - University of New South Wales agreement via the Council of Australian University Librarians.

### Biographies

Juan Ríos-Ocampo is a PhD candidate of the School of Management and Governance at UNSW Business School in Sydney, Australia. He holds a BS in Economics and ME in Systems Engineering from the National University of Colombia. Juan's research focuses on how stakeholders' mental models and decision-making of urban planning impact societal well-being in cities.

Michael Shayne Gary is a Professor of at UNSW Business School in Sydney Australia. He holds a BS in Management Science from the Massachusetts Institute of Technology (MIT) and a PhD in Strategy and International Management from London Business School. Shayne's research in behavioral strategy examines how differences in managerial mental models and implementation policies lead to differences in organizational performance.



## References

- Al-Eideh BM, Al-Omar HO. 2019. Population projection model using exponential growth function with a birth and death diffusion growth rate processes. *European Journal of Scientific Research* **151**(3): 271–276.
- Dennis B. 1989. Allee effects: population growth, critical density, and the chance of extinction. *Natural Resource Modeling* **3**(4): 481–538.
- Eker S, Slinger J, van Daalen E, Yücel G. 2014. Sensitivity analysis of graphical functions. *System Dynamics Review* **30**(3): 186–205. <https://doi.org/10.1002/sdr.1518>.
- Ford DN, Sterman JD. 1998. Expert knowledge elicitation to improve formal and mental models. *System Dynamics Review* **14**(4): 309–340.
- Forrester JW. 1961. *Industrial dynamics*. MIT Press: Cambridge, MA.
- Forrester JW. 1968. Market growth as influenced by capital investment. *Industrial Management Review (MIT) (currently published as the Sloan Management Review)* **9**(2): 83–105.
- Forrester JW. 1969. *Urban dynamics*. Productivity Press: Cambridge, MA.
- Forrester JW. 1971. *World dynamics*. Nippon Keiei Shuppankei: Tokyo.
- Gonçalves PM, Hines JH, Sterman JD. 2005. The impact of endogenous demand on push-pull production systems. *System Dynamics Review* **21**(3): 187–216.
- Lilien G, Rangaswamy A, de Bruyn A. 2017. *Principles of marketing engineering and analytics*, 3rd ed. DecisionPro, Inc: State College, PA.
- Martin, L. 1997. *Mistakes and misunderstandings: table functions*. Retrieved from Boston, USA.
- Saeed K. 1984. Continuous non-linear functions for use in system dynamics modelling. *DYNAMICA* **10**: 16–23.
- Sterman JD. 2000. *Business dynamics: systems thinking and modeling for a complex world*. Irwin/McGraw-Hill: New York.
- Struben J. 2020. The coronavirus disease (COVID-19) pandemic: simulation-based assessment of outbreak responses and postpeak strategies. *System Dynamics Review* **36**(3): 247–293. <https://doi.org/10.1002/sdr.1660>.
- Struben J, Sterman JD. 2008. Transition challenges for alternative fuel vehicle and transportation systems. *Environment and Planning B: Urban Analytics and City Science* **35**(6): 1070–1097. <https://doi.org/10.1068/b33022t>.

## Supporting information

Additional supporting information may be found in the online version of this article at the publisher's website.

## Appendix S1

### Data S1